

Summary

Which algorithm do we use in practice?

1. **Algorithms:**
 - a. GD
 - b. Momentum based GD
 - c. Nesterov Accelerated GD
 - d. AdaGrad
 - e. RMSProp
 - f. Adam
2. **Strategies:**
 - a. Batch
 - b. Mini-Batch (32, 64, 128)
 - c. Stochastic
3. Up till this point, we have been using GD with Batch update
4. In practice, Adam with Mini-Batch is the most popular choice
5. However, GD with Stochastic update is also used along with the use of special strategies to adjust the learning rate.